```
2020-08-04
------------------
Minu
----
    1. loop, if-else removal, map/filter
    2. sampled using suffle.
    Questions:
    ----------
    3.

Astha
------------------
    1. Vectorizer and angle calculation done.
    Questions:
    2.

Shankar
------------------
    1. Indexed vectorizer done.
    2. Environment conflicts.
    3. Performance evaluation.
------------------

Labin
------------------
palindome identification
------------------------
s1 = maaaam
s2 = madam

    >> def [x1, (reverse s1)]
    >> (= x1 s1)

[m, a, d, a, m ]

partition  = [m, a]
           = [a, m]

           = [m, a, a]
           = [a, a, m]

push => m, a, pop a, m

    (a, pop, a)
    (m, pop, m)
    not-equal => false.

     (= a a) each pop
     (= a a)
     (= m m)

-------------------
mid-position, start push, after, start pop.
-------------------
1M length.
-------------------
w1          w2
eat     vs  ate
feet    vs  foot

sim(w1, w2) = ???

edit distance
-------------
```

```
insertion
deletion
replacement
transposition

eat -> ate => 4 character
eat => delete(e)
at => insert(e)

edit distance =>

lab => labin : i, n insert

1M words
--------
a-z 26 characters ,

lab => labing
3 len(incomplete word) => 6 dictionary
3 to 6 operations choose from 26 characters.
-----------------------------------------
based on character replacement, .....

-----------------------------------------
eat => ate

you(person) eat rice(food).
he(person) ate fruit(food).
.........................................

Suggestion problem
----------------------
Words are iid. (indepent and identically distributed random variable)

And also,
the n-grams.
----------------------
Simplify the problem.


----------------------
n = 2, 3, 4, 5....
how do we choose it ?

shortest = 3
longest = 22

what happens
if n = 1, 2: over suggestion / over fitting

or n = 6, 7: under suggestion / under fitting
???

-2d <=> +2d
-----------------------------
n 2, 3, 4, 5, 6
-----------------------------
We use multi-grams
n = [2, 3], [3, 4], [4, 5], [5, 6]
n = [2, 3, 4], [3, 4, 5], [4, 5, 6]
-----------------------------
Parameter estimation.
-----------------------------
Test data
```

```
---------
word1 => suggestion1{}
word2 => suggestion2{}
.
.
.
wordn => suggestionn{}
------------------------
accuracy??, precision?? recall ??
we chosse best result of n.
???
------------------------
precision vs recall

if we increase precision, it will decrease recall and vice versa.
-----------------------------------------------------------------
Seen examples: Train
Unseen examples: Test


```